

SEQUENCE LISTING

SEQ ID NO: 1

SEQUENCE LENGTH: 4367

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION

TATCCGCACC TCCGCTCCCA CCCGGCGCCT CGGCGCGCCC GCCCTCCG ATG CGC TCA 57

Met Arg Ser

1

GCG GCC GCA GCT CCT CGG AGT CCC GCG GTG GCC ACC GAG TCT CGC CGC 105

Ala Ala Ala Ala Pro Arg Ser Pro Ala Val Ala Thr Glu Ser Arg Arg

5

10

15

TTC GCC GCA GCC AGG TGG CCC GGG TGG CGC TCG CTC CAG CGG CCG GCG 153

Phe Ala Ala Ala Arg Trp Pro Gly Trp Arg Ser Leu Gln Arg Pro Ala

20

25

30

35

CGG CGG AGC GGG CGG GGC GGC GGT GGC GCG GCC CCG GGA CCG TAT CCC 201

Arg Arg Ser Gly Arg Gly Gly Gly Gly Ala Ala Pro Gly Pro Tyr Pro

40

45

50

TCC GCC GCC CCT CCC CCG CCC GGC CCC GGC CCC CCT CCC TCC CGG CAG 249

Ser Ala Ala Pro Pro Pro Pro Gly Pro Gly Pro Pro Pro Ser Arg Gln

55

60

65

AGC TCG CCT CCC TCC GCC TCA GAC TGT TTT GGT AGC AAC GGC AAC GGC 297

Ser Ser Pro Pro Ser Ala Ser Asp Cys Phe Gly Ser Asn Gly Asn Gly

70

75

80

GGC GGC GCG TTT CGG CCC GGC TCC CGG CGG CTC CTT GGT CTC GGC GGC 345

Gly Gly Ala Phe Arg Pro Gly Ser Arg Arg Leu Leu Gly Leu Gly Gly

85	90	95	
CCT CCC CGC CCC TTC GTC GTC GTC CTT CTC CCC CTC GCC AGC CCG GGC			393
Pro Pro Arg Pro Phe Val Val Val Leu Leu Pro Leu Ala Ser Pro Gly			
100	105	110	115
GCC CCT CCG GCC GCG CCA ACC CGC GCC TCC CCG CTC GGC GCC CGT GCG			441
Ala Pro Pro Ala Ala Pro Thr Arg Ala Ser Pro Leu Gly Ala Arg Ala			
120	125	130	
TCC CCG CCG CGT TCC GGC GTC TCC TTG GCG CGC CCG GCT CCC GGC TGT			489
Ser Pro Pro Arg Ser Gly Val Ser Leu Ala Arg Pro Ala Pro Gly Cys			
135	140	145	
CCC CGC CCG GCG TGC GAG CCG GTG TAT GGG CCC CTC ACC ATG TCG CTG			537
Pro Arg Pro Ala Cys Glu Pro Val Tyr Gly Pro Leu Thr Met Ser Leu			
150	155	160	
AAG CCC CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAA			585
Lys Pro Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln			
165	170	175	
CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CCG CCG CCC GCG GCT GCC AAT			633
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Pro Pro Ala Ala Ala Asn			
180	185	190	195
GTC CGC AAG CCC GGC GGC AGC GGC CTT CTA GCG TCG CCC GCC GCC GCG			681
Val Arg Lys Pro Gly Gly Ser Gly Leu Leu Ala Ser Pro Ala Ala Ala			
200	205	210	
CCT TCG CCG TCC TCG TCC TCG GTC TCC TCG TCC TCG GCC ACG GCT CCC			729
Pro Ser Pro Ser Ser Ser Ser Val Ser Ser Ser Ser Ala Thr Ala Pro			
215	220	225	
TCC TCG GTG GTC GCG GCG ACC TCC GGC GGC GGC AGG CCC GGC CTG GGC			777
Ser Ser Val Val Ala Ala Thr Ser Gly Gly Gly Arg Pro Gly Leu Gly			
230	235	240	

AGA GGT CGA AAC AGT AAC AAA GGA CTG CCT CAG TCT ACG ATT TCT TTT	825
Arg Gly Arg Asn Ser Asn Lys Gly Leu Pro Gln Ser Thr Ile Ser Phe	
245 250 255	
GAT GGA ATC TAT GCA AAT ATG AGG ATG GTT CAT ATA CTT ACA TCA GTT	873
Asp Gly Ile Tyr Ala Asn Met Arg Met Val His Ile Leu Thr Ser Val	
260 265 270 275	
GTT GGC TCC AAA TGT GAA GTA CAA GTG AAA AAT GGA GGT ATA TAT GAA	921
Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn Gly Gly Ile Tyr Glu	
280 285 290	
GGA GTT TTT AAA ACT TAC AGT CCG AAG TGT GAT TTG GTA CTT GAT GCC	969
Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp Leu Val Leu Asp Ala	
295 300 305	
GCA CAT GAG AAA AGT ACA GAA TCC AGT TCG GGG CCG AAA CGT GAA GAA	1017
Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys Arg Glu Glu	
310 315 320	
ATA ATG GAG AGT ATT TTG TTC AAA TGT TCA GAC TTT GTT GTG GTA CAG	1065
Ile Met Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val Val Val Gln	
325 330 335	
TTT AAA GAT ATG GAC TCC AGT TAT GCA AAA AGA GAT GCT TTT ACT GAC	1113
Phe Lys Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala Phe Thr Asp	
340 345 350 355	
TCT GCT ATC AGT GCT AAA GTG AAT GGC GAA CAC AAA GAG AAG GAC CTG	1161
Ser Ala Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu Lys Asp Leu	
360 365 370	
GAG CCC TGG GAT GCA GGT GAA CTC ACA GCC AAT GAG GAA CTT GAG GCT	1209
Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu Leu Glu Ala	
375 380 385	
TTG GAA AAT GAC GTA TCT AAT GGA TGG GAT CCC AAT GAT ATG TTT CGA	1257

1257
 1209
 1161
 1113
 1065
 1017
 969
 921
 873
 825

Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp Met Phe Arg
 390 395 400
 TAT AAT GAA GAA AAT TAT GGT GTA GTG TCT ACG TAT GAT AGC AGT TTA 1305
 Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp Ser Ser Leu
 405 410 415
 TCT TCG TAT ACA GTG CCC TTA GAA AGA GAT AAC TCA GAA GAA TTT TTA 1353
 Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu Glu Phe Leu
 420 425 430 435
 AAA CGG GAA GCA AGG GCA AAC CAG TTA GCA GAA GAA ATT GAG TCA AGT 1401
 Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile Glu Ser Ser
 440 445 450
 GCC CAG TAC AAA GCT CGA GTG GCC CTG GAA AAC GAT GAT AGG AGT GAG 1449
 Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp Arg Ser Glu
 455 460 465
 GAA GAA AAA TAC ACA GCA GTT CAG AGA AAT TCC AGT GAA CGT GAG GGG 1497
 Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Ser Ser Glu Arg Glu Gly
 470 475 480
 CAC AGC ATA AAC ACT AGG GAA AAT AAA TAT ATT CCT CCT GGA CAA AGA 1545
 His Ser Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro Gly Gln Arg
 485 490 495
 AAT AGA GAA GTC ATA TCC TGG GGA AGT GGG AGA CAG AAT TCA CCG CGT 1593
 Asn Arg Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn Ser Pro Arg
 500 505 510 515
 ATG GGC CAG CCT GGA TCG GGC TCC ATG CCA TCA AGA TCC ACT TCT CAC 1641
 Met Gly Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser Thr Ser His
 520 525 530
 ACT TCA GAT TTC AAC CCG AAT TCT GGT TCA GAC CAA AGA GTA GTT AAT 1689
 Thr Ser Asp Phe Asn Pro Asn Ser Gly Ser Asp Gln Arg Val Val Asn

TCT	CCC	AGA	CAG	AAC	AGT	ATT	GGA	AAT	ACC	CCC	AGT	GGG	CCA	GTT	CTT	2169
Ser	Pro	Arg	Gln	Asn	Ser	Ile	Gly	Asn	Thr	Pro	Ser	Gly	Pro	Val	Leu	
695				700				705								
GCT	TCT	CCC	CAA	GCT	GGT	ATT	ATT	CCA	ACT	GAA	GCT	GTT	GCC	ATG	CCT	2217
Ala	Ser	Pro	Gln	Ala	Gly	Ile	Ile	Pro	Thr	Glu	Ala	Val	Ala	Met	Pro	
710				715				720								
ATT	CCA	GCT	GCA	TCT	CCT	ACG	CCT	GCT	AGT	CCT	GCA	TCG	AAC	AGA	GCT	2265
Ile	Pro	Ala	Ala	Ser	Pro	Thr	Pro	Ala	Ser	Pro	Ala	Ser	Asn	Arg	Ala	
725				730				735								
GTT	ACC	CCT	TCT	AGT	GAG	GCT	AAA	GAT	TCC	AGG	CTT	CAA	GAT	CAG	AGG	2313
Val	Thr	Pro	Ser	Ser	Glu	Ala	Lys	Asp	Ser	Arg	Leu	Gln	Asp	Gln	Arg	
740				745				750				755				
CAG	AAC	TCT	CCT	GCA	GGG	AAT	AAA	GAA	AAT	ATT	AAA	CCC	AAT	GAA	ACA	2361
Gln	Asn	Ser	Pro	Ala	Gly	Asn	Lys	Glu	Asn	Ile	Lys	Pro	Asn	Glu	Thr	
760				765				770								
TCA	CCT	AGC	TTC	TCA	AAA	GCT	GAA	AAC	AAA	GGT	ATA	TCA	CCA	GTT	GTT	2409
Ser	Pro	Ser	Phe	Ser	Lys	Ala	Glu	Asn	Lys	Gly	Ile	Ser	Pro	Val	Val	
775				780				785								
TCT	GAA	CAT	AGA	AAA	CAG	ATT	GAT	GAT	TTA	AAG	AAA	TTT	AAG	AAT	GAT	2457
Ser	Glu	His	Arg	Lys	Gln	Ile	Asp	Asp	Leu	Lys	Lys	Phe	Lys	Asn	Asp	
790				795				800								
TTT	AGG	TTA	CAG	CCA	AGT	TCT	ACT	TCT	GAA	TCT	ATG	GAT	CAA	CTA	CTA	2505
Phe	Arg	Leu	Gln	Pro	Ser	Ser	Thr	Ser	Glu	Ser	Met	Asp	Gln	Leu	Leu	
805				810				815								
AAC	AAA	AAT	AGA	GAG	GGA	GAA	AAA	TCA	AGA	GAT	TTG	ATC	AAA	GAC	AAA	2553
Asn	Lys	Asn	Arg	Glu	Gly	Glu	Lys	Ser	Arg	Asp	Leu	Ile	Lys	Asp	Lys	
820				825				830				835				
ATT	GAA	CCA	AGT	GCT	AAG	GAT	TCT	TTC	ATT	GAA	AAT	AGC	AGC	AGC	AAC	2601

Ile Glu Pro Ser Ala Lys Asp Ser Phe Ile Glu Asn Ser Ser Ser Asn	
840	845
850	
TGT ACC AGT GGC AGC AGC AAG CCG AAT AGC CCC AGC ATT TCC CCT TCA	2649
Cys Thr Ser Gly Ser Ser Lys Pro Asn Ser Pro Ser Ile Ser Pro Ser	
855	860
865	
ATA CTT AGT AAC ACG GAG CAC AAG AGG GGA CCT GAG GTC ACT TCC CAA	2697
Ile Leu Ser Asn Thr Glu His Lys Arg Gly Pro Glu Val Thr Ser Gln	
870	875
880	
GGG GTT CAG ACT TCC AGC CCA GCA TGT AAA CAA GAG AAA GAC GAT AAG	2745
Gly Val Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys Asp Asp Lys	
885	890
895	
GAA GAG AAG AAA GAC GCA GCT GAG CAA GTT AGG AAA TCA ACA TTG AAT	2793
Glu Glu Lys Lys Asp Ala Ala Glu Gln Val Arg Lys Ser Thr Leu Asn	
900	905
910	915
CCC AAT GCA AAG GAG TTC AAC CCA CGT TCC TTC TCT CAG CCA AAG CCT	2841
Pro Asn Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln Pro Lys Pro	
920	925
930	
TCT ACT ACC CCA ACT TCA CCT CGG CCT CAA GCA CAA CCT AGC CCA TCT	2889
Ser Thr Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro Ser Pro Ser	
935	940
945	
ATG GTG GGT CAT CAA CAG CCA ACT CCA GTT TAT ACT CAG CCT GTT TGT	2937
Met Val Gly His Gln Gln Pro Thr Pro Val Tyr Thr Gln Pro Val Cys	
950	955
960	
TTT GCA CCA AAT ATG ATG TAT CCA GTC CCA GTG AGC CCA GGC GTG CAA	2985
Phe Ala Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro Gly Val Gln	
965	970
975	
CCT TTA TAC CCA ATA CCT ATG ACG CCC ATG CCA GTG AAT CAA GCC AAG	3033
Pro Leu Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn Gln Ala Lys	

980	985	990	995	
ACA TAT AGA GCA GTA CCA AAT ATG CCC CAA CAG CGG CAA GAC CAG CAT				3081
Thr Tyr Arg Ala Val Pro Asn Met Pro Gln Gln Arg Gln Asp Gln His				
	1000	1005	1010	
CAT CAG AGT GCC ATG ATG CAC CCA GCG TCA GCA GCG GGC CCA CCG ATT				3129
His Gln Ser Ala Met Met His Pro Ala Ser Ala Ala Gly Pro Pro Ile				
	1015	1020	1025	
GCA GCC ACC CCA CCA GCT TAC TCC ACG CAA TAT GTT GCC TAC AGT CCT				3177
Ala Ala Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr Ser Pro				
	1030	1035	1040	
CAG CAG TTC CCA AAT CAG CCC CTT GTT CAG CAT GTG CCA CAT TAT CAG				3225
Gln Gln Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His Tyr Gln				
	1045	1050	1055	
TCT CAG CAT CCT CAT GTC TAT AGT CCT GTA ATA CAG GGT AAT GCT AGA				3273
Ser Gln His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn Ala Arg				
1060	1065	1070	1075	
ATG ATG GCA CCA CCA ACA CAC GCC CAG CCT GGT TTA GTA TCT TCT TCA				3321
Met Met Ala Pro Pro Thr His Ala Gln Pro Gly Leu Val Ser Ser Ser				
	1080	1085	1090	
GCA ACT CAG TAC GGG GCT CAT GAG CAG ACG CAT GCG ATG TAT GCA TGT				3369
Ala Thr Gln Tyr Gly Ala His Glu Gln Thr His Ala Met Tyr Ala Cys				
	1095	1100	1105	
CCC AAA TTA CCA TAC AAC AAG GAG ACA AGC CCT TCT TTC TAC TTT GCC				3417
Pro Lys Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr Phe Ala				
	1110	1115	1120	
ATT TCC ACG GGC TCC CTT GCT CAG CAG TAT GCG CAC CCT AAC GCT ACC				3465
Ile Ser Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn Ala Thr				
	1125	1130	1135	

CTG CAC CCA CAT ACT CCA CAC CCT CAG CCT TCA GCT ACC CCC ACT GGA 3513
 Leu His Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro Thr Gly
 1140 1145 1150 1155
 CAG CAG CAA AGC CAA CAT GGT GGA AGT CAT CCT GCA CCC AGT CCT GTT 3561
 Gln Gln Gln Ser Gln His Gly Gly Ser His Pro Ala Pro Ser Pro Val
 1160 1165 1170
 CAG CAC CAT CAG CAC CAG GCC GCC CAG GCT CTC CAT CTG GCC AGT CCA 3609
 Gln His His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala Ser Pro
 1175 1180 1185
 CAG CAG CAG TCA GCC ATT TAC CAC GCG GGG CTT GCG CCA ACT CCA CCC 3657
 Gln Gln Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr Pro Pro
 1190 1195 1200
 TCC ATG ACA CCT GCC TCC AAC ACG CAG TCG CCA CAG AAT AGT TTC CCA 3705
 Ser Met Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Asn Ser Phe Pro
 1205 1210 1215
 GCA GCA CAA CAG ACT GTC TTT ACG ATC CAT CCT TCT CAC GTT CAG CCG 3753
 Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val Gln Pro
 1220 1225 1230 1235
 GCG TAT ACC AAC CCA CCC CAC ATG GCC CAC GTA CCT CAG GCT CAT GTA 3801
 Ala Tyr Thr Asn Pro Pro His Met Ala His Val Pro Gln Ala His Val
 1240 1245 1250
 CAG TCA GGA ATG GTT CCT TCT CAT CCA ACT GCC CAT GCG CCA ATG ATG 3849
 Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro Met Met
 1255 1260 1265
 CTA ATG ACG ACA CAG CCA CCC GGC GGT CCC CAG GCC GCC CTC GCT CAA 3897
 Leu Met Thr Thr Gln Pro Pro Gly Gly Pro Gln Ala Ala Leu Ala Gln
 1270 1275 1280
 AGT GCA CTA CAG CCC ATT CCA GTC TCG ACA ACA GCG CAT TTC CCC TAT 3945

3513
 3561
 3609
 3657
 3705
 3753
 3801
 3849
 3897
 3945

Ser Ala Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe Pro Tyr

1285

1290

1295

ATG ACG CAC CCT TCA GTA CAA GCC CAC CAC CAA CAG CAG TTG 3987

Met Thr His Pro Ser Val Gln Ala His His Gln Gln Gln Leu

1300

1305

1310

TAAGGCTGCC CTGGAGGAAC CGAAAGGCCA AATCCCTCC TCCCTTCTAC TGCTTCTACC 4047

AACTGGAAGC ACAGAAACT AGAATTCAT TTATTTTGT TTTAAATAT ATATGTTGAT 4107

TTCTTGTAAC ATCCAATAGG AATGCTAACA GTTCACTTGC AGTGGAAGAT ACTTGGACCG 4167

AGTAGAGGCA TTTAGGAACT TGGGGGCTAT TCCATAATC CATATGCTGT TTCAGAGTCC 4227

CGCAGGTACC CCAGCTCTGC TTGCCGAAAC TGGAAGTTAT TTATTTTTTA ATAACCCTTG 4287

AAAGTCATGA ACACATCAGC TAGCAAAAGA AGTAACAAGA GTGATTCTTG CTGCTATTAC 4347

TGCTAAAAAA AAAAAAAAAA 4367

SEQ ID NO: 2

SEQUENCE LENGTH: 203

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

CACCACCAGC AACAGCAACA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG 60

CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG 120

CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG 180

CAGCATCAGC GAAACTCTGG GCC 203

SEQ ID NO: 3

SEQUENCE LENGTH: 20

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

CACCACCAGC AACAGCAACA

20

SEQ ID NO: 4

SEQUENCE LENGTH: 20

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

GGCCCAGAGT TTCGTGATG

20

SEQ ID NO: 5

SEQUENCE LENGTH: 165

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG 60

CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG 120

CAGCAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC AGCAG 165

SEQ ID NO: 6

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

CCCTCACCAT GTCGCTGAAG C

21

SEQ ID NO: 7

SEQUENCE LENGTH: 19

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

CGACGCTAGA AGGCCGCTG

19

SEQ ID NO: 8

SEQUENCE LENGTH: 19

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

CTTGCGGACA TTGGCAGCC

19

SEQ ID NO: 9

SEQUENCE LENGTH: 27

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

TTCTCTCAGC CAAAGCCTTC TACTACC

27

SEQ ID NO: 10

SEQUENCE LENGTH: 19

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

TATCCGCAGC TCCGCTCCG

19

SEQ ID NO: 11

SEQUENCE LENGTH: 20

SEQUENCE TYPE: nucleic acid

STRNDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION

AGCCGGGCCG AAACGCGCCG

20